

Analysis and Optimization of Interference Nulling in Downlink Multi-Antenna HetNets with Offloading

Yueping Wu*, Ying Cui[‡], Bruno Clerckx*[†]

*Department of Electrical and Electronic Engineering, Imperial College London [†]School of Electrical Engineering, Korea University

[‡]Department of Electronic Engineering, Shanghai Jiao Tong University

Abstract—Heterogeneous networks (HetNets) with offloading is considered as an effective way to meet the high data rate demand of future wireless service. However, the offloaded users suffer from strong inter-tier interference, which reduces the benefits of offloading and is one of the main limiting factors of the system performance. In this paper, we investigate the use of an interference nulling (IN) beamforming scheme to improve the system performance by carefully managing the inter-tier interference to the offloaded users in downlink two-tier HetNets with multi-antenna base stations. Utilizing tools from stochastic geometry, we derive a tractable expression for the rate coverage probability of the IN scheme. Then, we optimize the design parameter, i.e., the degrees of freedom that can be used for IN, to maximize the rate coverage probability. Specifically, in the asymptotic scenario where the rate threshold is small, by studying the order behavior of the rate coverage probability, we characterize the optimal design parameter. For the general scenario, we show some properties of the optimal design parameter. Finally, by numerical simulations, we show the IN scheme can outperform both the simple offloading scheme without interference management and the almost blank subframes scheme in 3GPP LTE, especially in large antenna regime.

I. INTRODUCTION

The next generation of wireless networks will see a significant increase in the number of wireless users and the scope of high data rate applications, leading to a high data rate demand. The conventional cellular solution, which comprises of high power base stations (BSs) covering a large cellular area, will not be able to scale with the increasing data rate demand of these future networks. A promising solution is the deployment of low power small cell nodes overlaid with high power macro-BSs, so called heterogeneous networks (HetNets), which are capable of aggressive reuse of existing spectrum assets, and thus lead to the support of future high data rate applications.

Due to the larger power at the macro-BSs, more users intend to associate with the macro-cell tier, which causes the problem of *load imbalance* between the macro-cell tier and the small-cell tier. Load imbalance is recognized as one of the main bottlenecks for the performance of HetNets [1, 2]. As such, to remit this problem, it is suggested to offload users from the heavily loaded macro-cell tier to the lightly loaded small-cell tier. Under the offloading scheme, the nearest macro-BS of each offloaded user, which provides the strongest signal among all the BSs, now becomes the dominant interferer of

this offloaded user. Thus, after offloading, the offloaded users have degraded signal-to-interference ratio (SIR). The performance of the offloaded users is one of the limiting factors of the network performance. Several interference management techniques have been considered to improve the performance of the offloaded users in HetNets with offloading. One such technique is almost blank subframes (ABS) in 3GPP LTE. In ABS, (time or frequency) resource is partitioned, whereby the offloaded users and the other users are served using different portions of the resource. The performance of ABS in HetNets with offloading was analyzed in [3] using tools from stochastic geometry. Another interference management technique was proposed in [4] to reduce the interference to offloaded users, where each offloaded user is jointly served by both of its nearest macro-BS and nearest pico-BS. The performance of this joint transmission scheme in large HetNets was analyzed using similar tools as in [3]. Note that these works only considered HetNets with single-antenna BSs.

Deploying multiple antennas in HetNets is another solution to provide high data rates for future wireless service. Furthermore, with multiple antennas, more effective interference management techniques (e.g., coordinated beamforming [5]) can be implemented. For example, a HetNet with a single multi-antenna macro-BS was investigated in [6, 7], where the multiple antennas at the macro-BS are used for serving its scheduled users as well as mitigating interference to the receivers in small cells using different interference coordination schemes. These schemes have been analyzed and shown to have performance improvement. However, since only one macro-BS is considered, the results obtained in [6, 7] can not reflect the macro-tier interference, and thus can not offer accurate insights for practical networks. In [8], interference coordination among a *fixed* number of neighboring BSs was investigated in downlink large multi-antenna HetNets. However, this scheme may not fully exploit the spatial properties of the interference in large HetNets. Moreover, offloading was not considered in [8]. So far, it is still not clear how the interference coordination schemes and the system parameters affect the performance of large multi-antenna HetNets with offloading.

In this paper, we consider offloading in downlink two-tier large stochastic HetNets with multi-antenna BSs, and investigate the use of interference nulling (IN) beamforming to improve the performance of the offloaded users. The scheme has a design parameter, which is the degree of freedom U

that can be used at each macro-BS for IN to its offloaded users. In particular, each macro-BS utilizes the low-complexity zero-forcing beamforming (ZFBF) precoder to suppress interference to at most U offloaded users as well as boost the desired signal to its scheduled user. Note that interference coordination using beamforming technique in large stochastic HetNets causes spatial dependency among macro and pico cells, which is quite difficult to analyze in general [7]. In this paper, by adopting appropriate approximations and utilizing tools from stochastic geometry, we first present a tractable expression for the rate coverage probability of the IN scheme. To our best knowledge, this is the first work analyzing the interference coordination technique in large stochastic multi-antenna HetNets with offloading. To further improve the rate coverage probability, we consider the optimization of the design parameter. Note that optimization problems in large HetNets with single-antenna BSs were investigated (see e.g., [9]). In [9], the objective function is relatively simple, and bounds of the objective function are utilized to obtain near-optimal solutions. The optimization problem in large multi-antenna HetNets we consider is an integer programming problem with a very complicated objective function. Hence, it is quite challenging to obtain the optimal solution. First, for the asymptotic scenario where the rate threshold is small, by studying the order behavior of the rate coverage probability, we prove that the optimal design parameter converges to a fixed value, which is only related to the antenna number difference between each macro-BS and each pico-BS. Next, for the general scenario, we show that the optimal design parameter also depends on other system parameters. Finally, by numerical simulations, we show the IN scheme can outperform both the simple offloading scheme without interference management and ABS in 3GPP LTE, especially in large antenna regime.

II. SYSTEM MODEL

A. Downlink Two-Tier Heterogeneous Networks

We consider a two-tier HetNet where a macro-cell tier is overlaid with a pico-cell tier, as shown in Fig. 1(a). The locations of the macro-cell BSs and the pico-cell BSs are spatially distributed as two independent Homogeneous Poisson point processes (PPPs) Φ_1 and Φ_2 with densities λ_1 and λ_2 , respectively. The locations of the users are also distributed as an independent homogeneous PPP Φ_u with density λ_u . Without loss of generality (w.l.o.g.), denote the macro-cell tier as the 1st tier and the pico-cell tier as the 2nd tier. We consider the downlink transmission. Each macro-BS has N_1 antennas with a total transmission power P_1 , each pico-BS has N_2 antennas with a total transmission power P_2 ($< P_1$), and each user is equipped with a single antenna. We consider both large-scale fading and small-scale fading. Specifically, due to large-scale fading, in the j th tier ($j = 1, 2$), transmitted signals with distance r are attenuated by a factor $\frac{1}{r^{\alpha_j}}$, where $\alpha_j > 2$ is the path loss exponent of the j th tier. For small-scale fading, we assume Rayleigh fading channels, i.e., each element of channel vectors is distributed as $\mathcal{CN}(0, 1)$.

B. User Association

We assume open access [1]. Due to the larger power at the macro-BSs, the load imbalancing problem arises if the user association is according to the long-term average received power (RP). To remit the load imbalancing problem, the bias factor B_j ($j = 1, 2$) is introduced to each tier, where $B_2 > B_1$, to offload users from the macro-cell tier to the pico-cell tier. Specifically, user i (denoted as u_i) is associated with the BS which provides the maximum *long-term averaged* biased-received-power (BRP) (among all the macro-BSs and pico-BSs) at the user. Here, the long-term averaged BRP is defined as the average RP multiplied by a bias factor. This associated BS is called the *serving BS* of user i . Note that within each tier, the nearest BS to user i provides the strongest long-term averaged BRP in this tier. User i is thus associated with the nearest BS in the j_i^* th tier if¹ $j_i^* = \arg \max_{j \in \{1, 2\}} P_j B_j Z_{i,j}^{-\alpha_j}$, where $Z_{i,j}$ is the distance between user i and its nearest BS in the j th tier. From the criterion described above, we observe that, for given $\{P_j\}$, $\{Z_{i,j}\}$ and $\{\alpha_j\}$, the user association is only affected by the bias factor ratio. Thus, w.l.o.g., we assume $B_1 = 1$ and $B_2 = B > 1$. After user association, each BS schedules its associated users according to TDMA, i.e., scheduling one user in each time slot, so that there is no intra-cell interference.

According to the above mentioned user association policy and the offloading strategy, as illustrated in Fig. 1(b), all the users can be partitioned into the following three disjoint user sets: i) set of *macro-users*: $\mathcal{U}_1 = \{u_i | P_1 Z_{i,1}^{-\alpha_1} \geq B P_2 Z_{i,2}^{-\alpha_2}\}$, ii) set of *unoffloaded pico-users*: $\mathcal{U}_{2\bar{O}} = \{u_i | P_2 Z_{i,2}^{-\alpha_2} > P_1 Z_{i,1}^{-\alpha_1}\}$, iii) set of *offloaded users*: $\mathcal{U}_{2O} = \{u_i | P_2 Z_{i,2}^{-\alpha_2} \leq P_1 Z_{i,1}^{-\alpha_1} < B P_2 Z_{i,2}^{-\alpha_2}\}$, where the macro-users are associated with the macro-BSs, the unoffloaded pico-users are associated with the pico-BSs (even without bias), and the offloaded users are offloaded from the macro-BSs to the pico-BSs (due to bias $B > 1$). Moreover, $\mathcal{U}_2 = \mathcal{U}_{2\bar{O}} \cup \mathcal{U}_{2O}$ is the set of *pico-users*.

C. Performance Metric

In this paper, we study the performance of the typical user denoted as² u_0 , which is located at the origin and is scheduled. Since HetNets are interference-limited, in this paper, we ignore the thermal noise in the analysis. Note that the analysis including thermal noise can be obtained in a similar way. Let $R_0 = \frac{W}{L_0} \log_2(1 + \text{SIR}_0)$ denote the rate of the typical user, where W is the available resource (e.g., time or frequency), L_0 is the total number of associated users (i.e., *load*) of the typical user's serving BS, and SIR_0 is the SIR of the typical user. We investigate the *rate coverage probability* of the typical user, which is defined as the probability that the rate of the typical user is larger than a threshold [3], i.e.,

$$\mathcal{R}(\tau) \triangleq \Pr(R_0 > \tau) = \Pr\left(\frac{W}{L_0} \log_2(1 + \text{SIR}_0) > \tau\right) \quad (1)$$

¹In the cell selection procedure, the first antenna is normally used to transmit signal (using the total transmission power of each BS) for BRP determination according to LTE standards.

²The index of the typical user and its serving BS is denoted as 0.

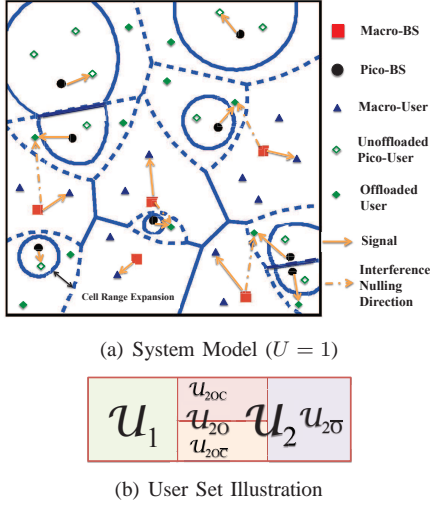


Fig. 1. System model ($U = 1$) and user set illustration.

where τ is the rate threshold. The rate coverage probability is a desired performance metric for applications with strict rate requirements, e.g., video services [2].

III. INTER-TIER INTERFERENCE NULLING

In HetNets with offloading, the offloaded users (to the pico-cell tier) normally suffer from stronger interference than the macro-users and unoffloaded pico-users.³ The dominant interference to each offloaded user, which is caused by its nearest macro-BS, is one of the limiting factors of the system performance. In this section, we first investigate an inter-tier IN scheme to improve the performance of these offloaded users. Then, we obtain the SIR under the IN scheme.

A. Strategy Description

We consider an inter-tier IN scheme to mitigate the dominant interference to the offloaded users. We first partition the offloaded users \mathcal{U}_{2O} into two sets. Note that all the offloaded users may not be scheduled by their nearest pico-BSs simultaneously, as each BS schedules one user in each time slot. We refer to the offloaded users, which are offloaded from a macro-BS and are scheduled by their nearest pico-BSs in a particular time slot, as *active offloaded users* of this macro-BS. In this scheme, each macro-BS conducts IN to some of its active offloaded users in a particular time slot, which are referred to as *IN offloaded users* of this macro-BS. We refer to the remaining offloaded users of each macro-BS as *non-IN offloaded users* of this macro-BS. Under the IN scheme, in a particular time slot, as illustrated in Fig. 1(b), the offloaded users \mathcal{U}_{2O} can be divided into two sets, i.e., $\mathcal{U}_{2O} = \mathcal{U}_{2OC} \cup \mathcal{U}_{2OC}$, where \mathcal{U}_{2OC} is the set of IN offloaded users and \mathcal{U}_{2OC} is the set of non-IN offloaded users.

We now discuss how to determine the IN offloaded users of each macro-BS. In particular, each macro-BS, which has

³For each offloaded user, its nearest macro-BS, which provides the strongest long-term averaged RP, now becomes the dominant interferer of this offloaded user. However, for each macro-user or unoffloaded pico-user, the BS which provides the strongest long-term averaged RP is its serving BS. Therefore, the offloaded users suffer from the strongest interference.

TABLE I
PARAMETER VALUES WHEN $u_0 \in \mathcal{U}_k$ WITH $k \in \{1, 2\bar{O}, 2OC, 2OC\}$

k	1	$2\bar{O}$	$2OC$	$2OC$
j_k	1	2	2	2
M_k	$N_1 - u_{2OC,0}$	N_2	N_2	N_2
\mathcal{B}_k	$\{B_{1,0}\}$	$\{B_{2,0}\}$	$\{B_{1,0}, B_{2,0}\}$	$\{B_{2,0}\}$

N_1 transmit antennas, makes use of at most U ($U < N_1$) DoF to conduct IN. Note that U is the design parameter of this scheme. When $U = 0$, the IN scheme reduces to the one without interference management technique. Specifically, let $U_{2Oa,\ell}$ denote the number of active offloaded users of macro-BS ℓ , each of which is scheduled by a different pico-BS. If $U_{2Oa,\ell} \leq U$, macro-BS ℓ performs IN to all of its $U_{2Oa,\ell}$ active offloaded users using $U_{2Oa,\ell}$ DoF. However, if $U_{2Oa,\ell} > U$, macro-BS ℓ randomly selects U out of $U_{2Oa,\ell}$ active offloaded users according to the uniform distribution to perform IN using U DoF. Hence, macro-BS ℓ performs IN to $u_{2OC,\ell} \triangleq \min(U, U_{2Oa,\ell})$ out of $U_{2Oa,\ell}$ active offloaded users. Note that the DoF used for IN (called IN DoF) at macro-BS ℓ is $u_{2OC,\ell}$. The remaining $N_1 - u_{2OC,\ell}$ DoF at macro-BS ℓ is used for boosting the signal to its scheduled user.

Now, we introduce the precoding vectors at macro-BSs and pico-BSs, respectively. First, each macro-BS utilizes the low-complexity ZFBF vector to serve its scheduled user and simultaneously perform IN to its IN offloaded users. Specifically, denote $\mathbf{H}_{1,\ell} = [\mathbf{h}_{1,\ell} \mathbf{g}_{1,\ell 1} \mathbf{g}_{1,\ell 2} \dots \mathbf{g}_{1,\ell u_{2OC,\ell}}]^H$, where⁴ $\mathbf{h}_{1,\ell} \stackrel{d}{\sim} \mathcal{CN}_{N_1,1}(\mathbf{0}_{N_1 \times 1}, \mathbf{I}_{N_1})$ is the channel vector between macro-BS ℓ and its scheduled user, and $\mathbf{g}_{1,\ell i} \stackrel{d}{\sim} \mathcal{CN}_{N_1,1}(\mathbf{0}_{N_1 \times 1}, \mathbf{I}_{N_1})$ denotes the channel vector between macro-BS ℓ and its IN offloaded user i ($i = 1, \dots, u_{2OC,\ell}$). The precoding matrix is designed to be $\mathbf{W}_{1,\ell} = \mathbf{H}_{1,\ell}^H (\mathbf{H}_{1,\ell} \mathbf{H}_{1,\ell}^H)^{-1}$. Then, the ZFBF vector at macro-BS ℓ is designed to be $\mathbf{f}_{1,\ell} = \frac{\mathbf{w}_{1,\ell}}{\|\mathbf{w}_{1,\ell}\|}$, where $\mathbf{w}_{1,\ell}$ is the first column of the precoding matrix $\mathbf{W}_{1,\ell}$. Next, the beamforming vector at each pico-BS is utilized to serve its scheduled user. Specifically, we use the maximum-ratio transmission scheme. The beamforming vector at pico-BS ℓ is $\mathbf{f}_{2,\ell} = \frac{\mathbf{h}_{2,\ell}}{\|\mathbf{h}_{2,\ell}\|}$, where $\mathbf{h}_{2,\ell} \stackrel{d}{\sim} \mathcal{CN}_{N_2,1}(\mathbf{0}_{N_2 \times 1}, \mathbf{I}_{N_2})$ is the channel vector between pico-BS ℓ and its scheduled user.

B. SIR of the Typical User

We now obtain the SIR expressions of the typical user $u_0 \in \mathcal{U}_k$, for all $k \in \mathcal{K} \triangleq \{1, 2\bar{O}, 2OC, 2OC\}$. Specifically, the SIR of the typical user $u_0 \in \mathcal{U}_k$ is given by⁵

$$\text{SIR}_{k,0} = \frac{\frac{\rho_{j_k}}{Y_{j_k}^{\alpha_{j_k}}} |\mathbf{h}_{j_k,0}^H \mathbf{f}_{j_k,0}|^2}{\sum_{j=1}^2 \sum_{\ell \in \Phi(\lambda_j) \setminus \mathcal{B}_k} \frac{\rho_j |\mathbf{h}_{j,\ell 0}^H \mathbf{f}_{j,\ell}|^2}{D_{j,\ell 0}^{\alpha_j}}}, \text{ where } \mathbf{h}_{j_k,0} \text{ is the}$$

channel vector between $u_0 \in \mathcal{U}_k$ and its serving BS, $\mathbf{f}_{j_k,0}$ is the beamforming vector at the serving BS of $u_0 \in \mathcal{U}_k$, with $|\mathbf{h}_{j_k,0}^H \mathbf{f}_{j_k,0}|^2 \stackrel{d}{\sim} \text{Gamma}(M_k, 1)$, j_k and M_k given by Table I, $\mathbf{h}_{j,\ell 0}$ is the channel vector between BS ℓ in the j th tier

⁴The notation $X \stackrel{d}{\sim} Y$ means that X is distributed as Y .

⁵In this paper, all macro-BSs and pico-BSs are assumed to be active. The same assumption can also be seen in existing papers (see e.g., [2, 3]).

and $u_0, \mathbf{f}_{j,\ell}$ is the beamforming vector at BS ℓ in the j th tier, with $\left| \mathbf{h}_{j,\ell}^H \mathbf{f}_{j,\ell} \right|^2 \stackrel{d}{\sim} \text{Gamma}(1, 1)$ ($j = 1, 2$), ρ_j is the transmit signal-to-noise ratio of the j th tier, Y_{jk} is the distance between $u_0 \in \mathcal{U}_k$ and its serving BS, $|D_{j,\ell_0}|$ is the distance between BS ℓ in the j th tier and u_0 , and \mathcal{B}_k (given by Table I) is the BS set which does not cause interference to $u_0 \in \mathcal{U}_k$. Here, $B_{j,0}$ is the nearest BS of u_0 in the j th tier.

IV. RATE COVERAGE PROBABILITY ANALYSIS

In this section, we investigate the rate coverage probability of the IN scheme. We notice that the rate coverage probability of the IN scheme is dependent on the probability mass functions (p.m.f.) of the active offloaded users $U_{2O_{a,0}}$ of the serving macro-BS of u_0 when $u_0 \in \mathcal{U}_1$ and the p.m.f. of the active offloaded users $\hat{U}_{2O_{a,0}}$ of the nearest macro-BS of u_0 when $u_0 \in \mathcal{U}_{2O}$. Note that the p.m.f.s of $U_{2O_{a,0}}$ and $\hat{U}_{2O_{a,0}}$ are related to Poisson distribution. But the exact distributions are unknown, as it is difficult to calculate due to the coupling between the active offloaded users and the pico-BSs. Here, we address this challenge by first deriving approximations for p.m.f.s of $U_{2O_{a,0}}$ and $\hat{U}_{2O_{a,0}}$. Based on the approximations and according to total probability theorem, we have the rate coverage probability $\mathcal{R}(\tau)$ as follows:⁶

Theorem 1: The rate coverage probability is

$$\mathcal{R}(\tau) = \mathcal{A}_1 \mathcal{R}_1(\tau) + \mathcal{A}_{2\bar{O}} \mathcal{R}_{2\bar{O}}(\tau) + \mathcal{A}_{2O} \Pr(\mathcal{E}_{2OC,0}) \mathcal{R}_{2OC}(\tau) + \mathcal{A}_{2O} (1 - \Pr(\mathcal{E}_{2OC,0})) \mathcal{R}_{2O\bar{C}}(\tau) \quad (2)$$

where $\mathcal{A}_k \triangleq \Pr(u_0 \in \mathcal{U}_k)$ ($k \in \{1, 2\bar{O}, 2O\}$) are given in [3], $\Pr(\mathcal{E}_{2OC,0})$ is the probability that $u_0 \in \mathcal{U}_{2O}$ is selected for IN: $\Pr(\mathcal{E}_{2OC,0}) = \frac{U \lambda_1 \mathcal{A}_2}{\lambda_2 \mathcal{A}_{2O}} \left(1 - \left(1 + \frac{\lambda_2 \mathcal{A}_{2O}}{3.5 \lambda_1 \mathcal{A}_2} \right)^{-3.5} \right) - \sum_{n=1}^U \frac{U}{n} \Pr(\hat{U}_{2O_{a,0}} = n) + \sum_{n=1}^U \Pr(\hat{U}_{2O_{a,0}} = n)$, with $\mathcal{A}_2 = \mathcal{A}_{2O} + \mathcal{A}_{2\bar{O}}$, $\Pr(\hat{U}_{2O_{a,0}} = n) \approx \frac{3.5^{3.5} \Gamma(n+3.5)}{\Gamma(n) \Gamma(3.5)} \left(\frac{\lambda_2 \mathcal{A}_{2O}}{\mathcal{A}_2 \lambda_1} \right)^{n-1} \left(3.5 + \frac{\lambda_2 \mathcal{A}_{2O}}{\mathcal{A}_2 \lambda_1} \right)^{-(n+3.5)}$ ($n \geq 1$), $\mathcal{R}_k(\tau) \triangleq \mathbb{E}_{L_{0,jk}} \left[\mathcal{S}_k \left(f \left(\frac{L_{0,jk} \tau}{W} \right) \right) \right]$ for all $k \in \mathcal{K}$, where $L_{0,jk}$ is the load of the typical user's serving BS when $u_0 \in \mathcal{U}_k$. Specifically,

$$\mathcal{R}_1(\tau) = \sum_{n \geq 1} \Pr(L_{0,1} = n) \mathcal{S}_1 \left(f \left(\frac{n\tau}{W} \right) \right) \quad (3)$$

$$\mathcal{R}_{2\bar{O}}(\tau) = \sum_{n \geq 1} \Pr(L_{0,2} = n) \mathcal{S}_{2\bar{O}} \left(f \left(\frac{n\tau}{W} \right) \right) \quad (4)$$

$$\mathcal{R}_{2OC}(\tau) = \sum_{n \geq 1} \Pr(L_{0,2} = n) \mathcal{S}_{2OC} \left(f \left(\frac{n\tau}{W} \right) \right) \quad (5)$$

$$\mathcal{R}_{2O\bar{C}}(\tau) = \sum_{n \geq 1} \Pr(L_{0,2} = n) \mathcal{S}_{2O\bar{C}} \left(f \left(\frac{n\tau}{W} \right) \right) \quad (6)$$

where $\mathcal{S}_k(\cdot)$ are given by (7)–(10) at the top of the next page.

$f(x) = 2^x - 1$, $\Pr(L_{0,j} = n) = \frac{3.5^{3.5} \Gamma(n+3.5) \left(\frac{\lambda_u \mathcal{A}_j}{\lambda_j} \right)^{n-1}}{\Gamma(3.5)(n-1)! \left(\frac{\lambda_u \mathcal{A}_j}{\lambda_j} + 3.5 \right)^{n+3.5}}$ ($j = 1, 2$), $f_{Y_1}(y)$ and $f_{Y_2}(y)$ are given in [3], $f_{Y_1, Y_2}(x, y)$ is

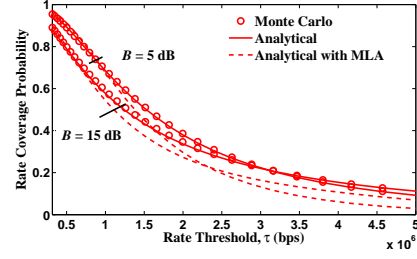


Fig. 2. Rate coverage probability vs. rate threshold τ for different bias factors B , and with $\alpha_1 = \alpha_2 = 4$, $\frac{P_1}{P_2} = 10$ dB, $N_1 = 8$, $N_2 = 4$, $U = 4$, $W = 10$ MHz, $\lambda_1 = 0.0001$ nodes/m², and $\lambda_2 = 0.0005$ nodes/m².

given in [4], $\mathcal{Q}_3 \triangleq \{(q_a)_{a=1}^3 | q_a \in \mathbb{N}^0, \sum_{a=1}^3 q_a = n\}$,

$$\tilde{\mathcal{L}}_{I_j}^{(m)}(s, r_{jk}) = \mathcal{L}_{I_j}(s, r_{jk}) \sum_{(p_a)_{a=1}^m \in \mathcal{M}_m} \frac{m!}{\prod_{a=1}^m p_a!} \times \prod_{a=1}^m \left(\frac{2\pi}{\alpha_j} \lambda_j s^{\frac{2}{\alpha_j}} B' \left(1 + \frac{2}{\alpha_j}, a - \frac{2}{\alpha_j}, \frac{1}{1 + sr_{jk}^{-\alpha_j}} \right) \right)^{p_a} \quad (11)$$

with $\mathcal{M}_m \triangleq \{(p_a)_{a=1}^m | p_a \in \mathbb{N}^0, \sum_{a=1}^m p_a = m\}$, $\mathcal{L}_{I_j}(s, r_{jk}) = \exp \left(-\frac{2\pi \lambda_j}{\alpha_j} s^{\frac{2}{\alpha_j}} B' \left(\frac{2}{\alpha_j}, 1 - \frac{2}{\alpha_j}, \frac{1}{1 + sr_{jk}^{-\alpha_j}} \right) \right)$, and $B'(a, b, z) = \int_z^1 u^{a-1} (1-u)^{b-1} du$ ($0 < z < 1$),

$$\Pr(u_{2OC,0} = n) = \begin{cases} \Pr(U_{2O_{a,0}} = n), & \text{for } 0 \leq n < U \\ \sum_{u=n}^{\infty} \Pr(U_{2O_{a,0}} = u), & \text{for } n = U \end{cases}$$

where $\Pr(U_{2O_{a,0}} = n) \approx \frac{3.5^{3.5} \Gamma(n+3.5)}{\Gamma(3.5)n!} \left(\frac{\lambda_2 \mathcal{A}_{2O}}{\mathcal{A}_2 \lambda_1} \right)^n \times \left(3.5 + \frac{\lambda_2 \mathcal{A}_{2O}}{\mathcal{A}_2 \lambda_1} \right)^{-(n+3.5)}$ ($n \geq 0$).

Note that the expression in (2) is difficult to compute and analyze due to the infinite summations over n . To simplify the rate coverage probability expression $\mathcal{R}(\tau)$ in (2), we use the mean of the random load (i.e., $\mathbb{E}[L_{0,1}]$ or $\mathbb{E}[L_{0,2}]$) to approximate it (i.e., $L_{0,1}$ or $L_{0,2}$) [3]. The simplification is thus achieved due to the elimination of the infinite summation over n . The rate coverage probability with mean load approximation (MLA) is given by:

Corollary 1: The rate coverage probability with MLA is

$$\bar{\mathcal{R}}(\tau) = \mathcal{A}_1 \bar{\mathcal{R}}_1(\tau) + \mathcal{A}_{2\bar{O}} \bar{\mathcal{R}}_{2\bar{O}}(\tau) + \mathcal{A}_{2O} \Pr(\mathcal{E}_{2OC,0}) \bar{\mathcal{R}}_{2OC}(\tau) + \mathcal{A}_{2O} (1 - \Pr(\mathcal{E}_{2OC,0})) \bar{\mathcal{R}}_{2O\bar{C}}(\tau) \quad (12)$$

where $\bar{\mathcal{R}}_1(\tau) = \mathcal{S}_1 \left(f \left(\frac{\mathbb{E}[L_{0,1}] \tau}{W} \right) \right)$, $\bar{\mathcal{R}}_{2\bar{O}}(\tau) = \mathcal{S}_{2\bar{O}} \left(f \left(\frac{\mathbb{E}[L_{0,2}] \tau}{W} \right) \right)$, $\bar{\mathcal{R}}_{2OC}(\tau) = \mathcal{S}_{2OC} \left(f \left(\frac{\mathbb{E}[L_{0,2}] \tau}{W} \right) \right)$, and $\bar{\mathcal{R}}_{2O\bar{C}}(\tau) = \mathcal{S}_{2O\bar{C}} \left(f \left(\frac{\mathbb{E}[L_{0,2}] \tau}{W} \right) \right)$ with $\mathcal{S}_k(\cdot)$ given by (7)–(10). Here, $\mathbb{E}[L_{0,1}] = 1 + 1.28 \frac{\lambda_u \mathcal{A}_1}{\lambda_1}$ [3], and $\mathbb{E}[L_{0,2}] = 1 + 1.28 \frac{\lambda_u \mathcal{A}_2}{\lambda_2}$.

Fig. 2 plots the rate coverage probability vs. rate threshold τ for different B . We see that the ‘Analytical’ curves (i.e., $\mathcal{R}(\tau)$ in (2)) closely match with the ‘Monte Carlo’ curves, although $\mathcal{R}(\tau)$ is derived based on the p.m.f. approximations of $U_{2O_{a,0}}$ and $\hat{U}_{2O_{a,0}}$. Moreover, we observe that the rate coverage probability with MLA (i.e., $\bar{\mathcal{R}}(\tau)$ in (12)) provides sufficient accuracy, especially when τ is not very large. Hence,

⁶Due to page limit, we omit all the proofs. Please refer to [10] for details.

$$\mathcal{S}_1(\beta) = \sum_{u=0}^U \left(\int_0^\infty \sum_{n=0}^{N_1-u-1} \frac{1}{n!} \sum_{n_1=0}^n \binom{n}{n_1} \tilde{\mathcal{L}}_{I_1}^{(n_1)}(s, y) \Big|_{s=\beta y^{\alpha_1}} \tilde{\mathcal{L}}_{I_2}^{(n-n_1)} \left(s, \left(\frac{P_2 B}{P_1} \right)^{\frac{1}{\alpha_2}} y^{\frac{\alpha_1}{\alpha_2}} \right) \Big|_{s=\beta y^{\alpha_1} \frac{\rho_2}{\rho_1}} f_{Y_1}(y) dy \right) \times \Pr(u_{2OC,0} = u) \quad (7)$$

$$\mathcal{S}_{2\bar{O}}(\beta) = \int_0^\infty \sum_{n=0}^{N_2-1} \frac{1}{n!} \sum_{n_1=0}^n \binom{n}{n_1} \tilde{\mathcal{L}}_{I_1}^{(n_1)} \left(s, \left(\frac{P_1}{P_2} \right)^{\frac{1}{\alpha_1}} y^{\frac{\alpha_2}{\alpha_1}} \right) \Big|_{s=\beta y^{\alpha_2} \frac{\rho_1}{\rho_2}} \tilde{\mathcal{L}}_{I_2}^{(n-n_1)}(s, y) \Big|_{s=\beta y^{\alpha_2}} f_{Y_2}(y) dy \quad (8)$$

$$\mathcal{S}_{2OC}(\beta) = \int_0^\infty \int_{\left(\frac{P_2}{P_1}\right)^{\frac{1}{\alpha_2}} x^{\frac{\alpha_1}{\alpha_2}}}^{\left(\frac{BP_2}{P_1}\right)^{\frac{1}{\alpha_2}} x^{\frac{\alpha_1}{\alpha_2}}} \sum_{n=0}^{N_2-1} \frac{1}{n!} \sum_{n_1=0}^n \binom{n}{n_1} \tilde{\mathcal{L}}_{I_1}^{(n_1)}(s, r_{1k}) \Big|_{s=\beta Y_{j_k}^{\alpha_{j_k}} \frac{\rho_1}{\rho_{j_k}}} \tilde{\mathcal{L}}_{I_2}^{(n-n_1)}(s, r_{2k}) \Big|_{s=\beta Y_{j_k}^{\alpha_{j_k}} \frac{\rho_2}{\rho_{j_k}}} f_{Y_1, Y_2}(x, y) dy dx \quad (9)$$

$$\mathcal{S}_{2O\bar{C}}(\beta) = \int_0^\infty \int_{\left(\frac{P_2}{P_1}\right)^{\frac{1}{\alpha_2}} x^{\frac{\alpha_1}{\alpha_2}}}^{\left(\frac{BP_2}{P_1}\right)^{\frac{1}{\alpha_2}} x^{\frac{\alpha_1}{\alpha_2}}} \sum_{n=0}^{N_2-1} \frac{1}{n!} \sum_{(q_a)_{a=1}^3 \in \mathcal{Q}_3} \binom{n}{q_1, q_2, q_3} \tilde{\mathcal{L}}_{I_1}^{(q_1)}(s, x) \Big|_{s=\beta y^{\alpha_2} \frac{\rho_1}{\rho_2}} \tilde{\mathcal{L}}_{I_2}^{(q_2)}(s, y) \Big|_{s=\beta y^{\alpha_2}} \Gamma(q_3 + 1) \times \left(\beta \frac{\rho_1 y^{\alpha_2}}{\rho_2 x^{\alpha_1}} \right)^{q_3} \left(1 + \beta \frac{\rho_1 y^{\alpha_2}}{\rho_2 x^{\alpha_1}} \right)^{-(q_3+1)} f_{Y_1, Y_2}(x, y) dy dx \quad (10)$$

for analytical tractability, we will investigate the rate coverage probability with MLA $\bar{\mathcal{R}}(\tau)$ in the remaining part of this paper.

V. RATE COVERAGE PROBABILITY OPTIMIZATION

In this section, we present some properties of the optimal design parameter which maximizes the rate coverage probability for small rate threshold regime and general rate threshold regime, respectively.

For a fixed B , the optimal design parameter $U^*(\tau)$ is defined as follows:⁷

$$U^*(\tau) \triangleq \arg \max_{U \in \{0, 1, \dots, N_1-1\}} \bar{\mathcal{R}}(U, \tau). \quad (13)$$

Note that (13) is an integer programming problem. Moreover, the objective function $\bar{\mathcal{R}}(U, \tau)$ in (12) is very complicated. It is thus quite challenging to obtain the closed-form optimal solution to the problem in (13).

Now, we address this challenge and characterize the optimality property of $U^*(\tau)$. Let $\Delta \bar{\mathcal{R}}(U, \tau) \triangleq \bar{\mathcal{R}}(U, \tau) - \bar{\mathcal{R}}(U-1, \tau)$ denote the rate coverage probability change when the design parameter is changed from $U-1$ to U . We can show

$$\Delta \bar{\mathcal{R}}(U, \tau) = \mathcal{A}_{2O} \Delta \bar{\mathcal{R}}_{2O}(U, \tau) - \mathcal{A}_1 |\Delta \bar{\mathcal{R}}_1(U, \tau)| \quad (14)$$

where $\Delta \bar{\mathcal{R}}_1(U, \tau) \triangleq \bar{\mathcal{R}}_1(U, \tau) - \bar{\mathcal{R}}_1(U-1, \tau) < 0$ denotes the rate coverage probability change of a macro-user, and $\Delta \bar{\mathcal{R}}_{2O}(U, \tau) \triangleq \bar{\mathcal{R}}_{2O}(U, \tau) - \bar{\mathcal{R}}_{2O}(U-1, \tau) > 0$ denotes the rate coverage probability change of an off-loaded user. Here⁸, $\bar{\mathcal{R}}_{2O}(U, \tau) = \Pr(U, \mathcal{E}_{2OC,0}) \bar{\mathcal{R}}_{2OC}(\tau) + (1 - \Pr(U, \mathcal{E}_{2OC,0})) \bar{\mathcal{R}}_{2O\bar{C}}(\tau)$. We observe from (14) that when U is increased by 1, $\Delta \bar{\mathcal{R}}(U, \tau)$ can be decomposed into two parts, namely, the “gain” $\mathcal{A}_{2O} \Delta \bar{\mathcal{R}}_{2O}(U, \tau)$ and the “penalty” $\mathcal{A}_1 |\Delta \bar{\mathcal{R}}_1(U, \tau)|$. Whether $\Delta \bar{\mathcal{R}}(U, \tau)$ is positive or not depends on whether the “gain” dominates the “penalty” or not. In the following, to maximize $\bar{\mathcal{R}}(U, \tau)$, we study

the properties of $\Delta \bar{\mathcal{R}}(U, \tau)$ in (14) w.r.t. U by comparing $\mathcal{A}_{2O} \Delta \bar{\mathcal{R}}_{2O}(U, \tau)$ and $\mathcal{A}_1 |\Delta \bar{\mathcal{R}}_1(U, \tau)|$.

1) *Rate coverage probability optimization when $\tau \rightarrow 0$:* We first characterize the asymptotic behavior of $\Delta \bar{\mathcal{R}}_{2O}(U, \tau)$ and $|\Delta \bar{\mathcal{R}}_1(U, \tau)|$ when $\tau \rightarrow 0$, which is shown as follows:

Lemma 1: When $\tau \rightarrow 0$, we have⁹ $\Delta \bar{\mathcal{R}}_{2O}(U, \tau) = \Theta(\tau^{N_2})$ and $|\Delta \bar{\mathcal{R}}_1(U, \tau)| = \Theta(\tau^{N_1-U})$.

According to (14), *Lemma 1*, and noting that \mathcal{A}_{2O} and \mathcal{A}_1 are independent of τ , we have

$$\Delta \bar{\mathcal{R}}(U, \tau) = \begin{cases} \Theta(\tau^{N_2}) > 0, & U < N_1 - N_2 \\ \Theta(\tau^{N_2}) - \Theta(\tau^{N_2}), & U = N_1 - N_2 \\ \Theta(\tau^{N_1-U}) < 0, & U > N_1 - N_2 \end{cases}. \quad (15)$$

Since $U^*(\tau)$ satisfies $\Delta \bar{\mathcal{R}}(U^*(\tau), \tau) > 0$ and $\Delta \bar{\mathcal{R}}(U^*(\tau) + 1, \tau) \leq 0$, we see from (15) that $U^*(\tau)$ should be in the set $\{N_1 - N_2 - 1, N_1 - N_2\}$, and the exact value of $U^*(\tau)$ depends on whether $\Delta \bar{\mathcal{R}}(U, \tau)$ is positive or not when $U = N_1 - N_2$ (i.e., the second case in (15)), i.e., whether the coefficient in $\Theta(\tau^{N_2})$ of $\mathcal{A}_{2O} \Delta \bar{\mathcal{R}}_{2O}(U, \tau)$ (i.e., the first one) is larger than that in $\Theta(\tau^{N_2})$ of $\mathcal{A}_1 |\Delta \bar{\mathcal{R}}_1(U, \tau)|$ (i.e., the second one) or not. According to the above discussions, the optimal design parameter $U^*(\tau)$ is given in the following theorem:

Theorem 2: When $\tau \rightarrow 0$, the optimal design parameter $U^*(\tau) \rightarrow U_0^*$, where $U_0^* \in \{N_1 - N_2 - 1, N_1 - N_2\}$.

Theorem 2 shows that when $\tau \rightarrow 0$, the optimal design parameter $U^*(\tau)$ converges to a fixed value in the set $\{N_1 - N_2 - 1, N_1 - N_2\}$, which is only related to the number of antennas at each macro-BS and each pico-BS. Fig. 3(a) plots the optimal design parameter $U^*(\tau)$ vs. rate threshold τ for different B . We can see that $U^*(\tau)$ converges to a fixed value $U_0^* \in \{N_1 - N_2 - 1, N_1 - N_2\}$ when τ is sufficiently small. Specifically, we can see that $U^*(\tau) = N_1 - N_2 = 3$ at $B = 4.6$ dB, and $U^*(\tau) = N_1 - N_2 - 1 = 2$ at $B = 2.5$ dB. These observations verify *Theorem 2*. In addition, we see that $U^*(\tau)$

⁷We make explicit the dependence of $\bar{\mathcal{R}}(\tau)$ on U in this section.

⁸Here, we make explicit the dependence of $\Pr(\mathcal{E}_{2OC,0})$ on U .

⁹ $f(x) = \Theta(g(x))$ means that $\lim_{x \rightarrow 0} \frac{f(x)}{g(x)} = c$ where $0 < c < \infty$.

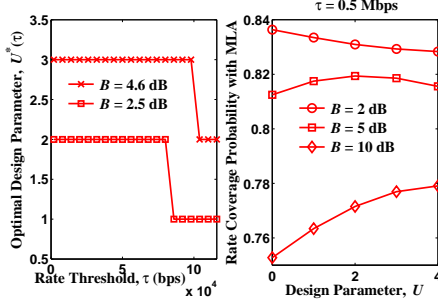
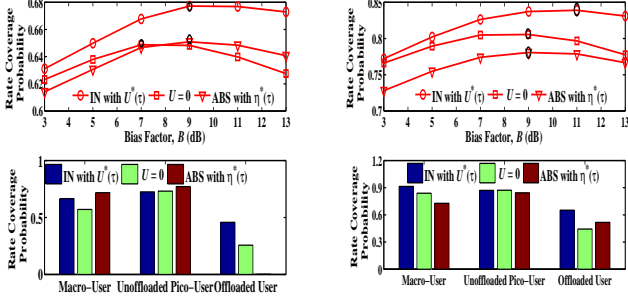


Fig. 3. Optimal design parameter $U^*(\tau)$, and with $\alpha_1 = \alpha_2 = 3$, $\frac{P_1}{P_2} = 10$ dB, $W = 10 \times 10^6$ Hz, $N_1 = 5$, $N_2 = 2$, $\lambda_u = 0.01$ nodes/m², $\lambda_1 = 0.0001$ nodes/m², and $\lambda_2 = 0.0015$ nodes/m².



(a) $N_1 = 8$, $N_2 = 6$, $\eta^*(\tau) = 0.01$ at B_{ABS}^* (b) $N_1 = 18$, $N_2 = 16$, $\eta^*(\tau) = 0.19$ at B_{ABS}^*

Fig. 4. Rate coverage probability vs. bias factors B , and with $\alpha_1 = 4.5$, $\alpha_2 = 4.7$, $\frac{P_1}{P_2} = 13$ dB, $W = 10 \times 10^6$ Hz, $\tau = 5 \times 10^5$ bps, $\lambda_1 = 0.00008$ nodes/m², $\lambda_2 = 0.001$ nodes/m², and $\lambda_u = 0.05$ nodes/m². In the figures on top, the points at B_{IN}^* , $B_{U=0}^*$, and B_{ABS}^* are highlighted using black ellipse. In the figures at the bottom, the rate coverage probability of each user type in different schemes are plotted at B_{IN}^* , $B_{U=0}^*$, and B_{ABS}^* , respectively.

is larger for a larger B .

2) *Rate coverage probability for general τ* : Unlike the case for small τ , the “gain” $\mathcal{A}_{2O} \Delta \bar{\mathcal{R}}_{2O}(U, \tau)$ and the “penalty” $\mathcal{A}_1 |\Delta \bar{\mathcal{R}}_1(U, \tau)|$ are not determined by the order terms $\Theta(\tau^{N_2})$ and $\Theta(\tau^{N_1-U})$, respectively. The optimal design parameter $U^*(\tau)$ thus also depends on other system parameters besides N_1 and N_2 . Fig. 3(b) plots the rate coverage probability with MLA vs. U for different B . We can see that besides $N_1 - N_2 - 1$ and $N_1 - N_2$, the optimal design parameter $U^*(\tau)$ can also take other values in set $\{0, 1, \dots, N_1 - 1\}$. In particular, from Fig. 3(b), we see that $U^*(\tau)$ can be 0 (at $B = 2$ dB), 2 (at $B = 5$ dB), and $N_1 - 1 = 4$ (at $B = 10$ dB). Interestingly, Fig. 3(b) indicates that, for general τ , the optimal parameter $U^*(\tau)$ increases with B , which is consistent with the case for small τ in Fig. 3(a).

VI. NUMERICAL RESULTS

In this section, we compare the *IN scheme* under $U^*(\tau)$ with the *simple offloading scheme without interference management* (i.e., $U = 0$) and *multi-antenna ABS* under $\eta^*(\tau)$ (i.e., the multi-antenna version of ABS in [3]). Specifically, $\eta \in (0, 1)$ is the continuous control variable for ABS, where η represents the resource fraction for serving the offloaded users only, while $1 - \eta$ represents the resource fraction for serving the macro-users and unoffloaded pico-users simultaneously. In the

simulation, the optimal control variable $\eta^*(\tau)$ of ABS is obtained by bisection method with N_1 iterations, while the optimal (discrete) design parameter $U^*(\tau)$ of the IN scheme is obtained by exhaustive search over $\{0, 1, \dots, N_1 - 1\}$.

Fig. 4 plots the rate coverage probability vs. bias factor B for all the three schemes. Note that under the parameters used in the simulation, we have sufficient offloaded users, and the dominant macro-interferer is sufficiently strong compared to the remaining macro-interferers. We see from Fig. 4 that the optimal rate coverage probability (maximized over B) of the IN scheme is larger than those of both the simple offloading scheme and ABS. ¹⁰ We denote the optimal B for the IN scheme, simple offloading scheme, and ABS as B_{IN}^* , $B_{U=0}^*$, and B_{ABS}^* , respectively. In particular, when B_{IN}^* is sufficiently large (i.e., sufficient offloaded users), if N_1 is sufficiently large (e.g., $N_1 = 8$), the IN scheme can achieve good performance gains over both the simple offloading scheme and ABS, as the penalty of sacrificing some DoF for IN becomes minor. From the figures at the bottom of Fig. 4(a) and Fig. 4(b), we see that the rate coverage probability of the offloaded users in the IN scheme at B_{IN}^* is larger than those in the simple offloading scheme at $B_{U=0}^*$ and in ABS at B_{ABS}^* . This is because the offloaded users in the IN scheme do not have (time or frequency) resource sacrifice, and their dominant macro-interferers can be cancelled. However, the offloaded users in ABS suffer from resource limitations, and the offloaded users in the simple offloading scheme suffer from the strong interference caused by their dominant macro-interferers.

REFERENCES

- [1] H. S. Jo, Y. J. Sang, P. Xia, and J. G. Andrews, “Heterogeneous cellular networks with flexible cell association: A comprehensive downlink SINR analysis,” *IEEE Trans. Wireless Commun.*, vol. 11, no. 10, pp. 3484–3495, Oct. 2012.
- [2] S. Singh, H. S. Dhillon, and J. G. Andrews, “Offloading in heterogeneous networks: modeling, analysis, and design insights,” *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 2484–2497, Mar. 2013.
- [3] S. Singh and J. G. Andrews, “Joint resource partitioning and offloading in heterogeneous cellular networks,” *IEEE Trans. Wireless Commun.*, vol. 13, no. 2, pp. 888–901, Feb. 2014.
- [4] A. H. Sakr and E. Hossain, “Location-aware cross-tier coordinated multipoint transmission in two-tier cellular networks,” *to appear in IEEE Trans. Wireless. Commun.*
- [5] Y. Cui, V. K. N. Lau, and Y. Wu, “Delay-aware BS discontinuous transmission control and user scheduling for energy harvesting downlink coordinated MIMO systems,” *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3786–3795, Jul. 2012.
- [6] K. Hosseini, J. Hoydis, S. t. Brink, and M. Debbah, “Massive MIMO and small cells: How to densify heterogeneous networks,” in *Proc. of IEEE Int. Conf. on Commun. (ICC)*, Budapest, Jun. 2013, pp. 5442–5447.
- [7] A. Adhikary, H. S. Dhillon, and G. Caire, “Massive-MIMO meets HetNet: Interference coordination through spatial blanking,” *submitted to IEEE J. Select. Areas Commun.*
- [8] P. Xia, C. H. Liu, and J. G. Andrews, “Downlink coordinated multipoint with overhead modeling in heterogeneous cellular networks,” *IEEE Trans. Wireless Commun.*, vol. 12, no. 8, pp. 4025–4037, Aug. 2013.
- [9] Y. Lin and W. Yu, “Joint spectrum partition and user association in multi-tier heterogeneous networks,” in *Proc. of Conference on Information Science and Systems (CISS)*, Princeton, NJ, Mar. 2014, pp. 1–6.
- [10] Y. Wu, Y. Cui, and B. Clerckx, “Analysis and optimization of inter-tier interference coordination in downlink multi-antenna HetNets with offloading,” 2014. [Online]. Available: <http://arxiv.org/abs/1411.3271>

¹⁰Note that the IN scheme may not provide gains in all scenarios, as suggested in Fig. 3.